USARIEM TECHNICAL NOTE

ENVIRONMENTAL MEDICINE GENOME BANK (EMGB): ANNUAL REPORT AND PROJECT SUMMARY

Nicholas J. Messinese[2]
Craig M. Lilly[2]
LTC Larry A. Sonna[1, 2]

[1]Thermal and Mountain Medicine Division
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760-5007

[2]Division of Pulmonary and Critical Care Medicine
Brigham and Women's Hospital
75 Francis Street
Boston, MA 02115

July 2003

U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760-5007

DISCLAIMERS

The views, opinions and/or assertions contained in this publication are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

Human subjects participated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR70-25 and USAMRMC Regulation 70-25 on the Use of Volunteers in Research. For protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law CFR 46.

Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2003	USARIEM Technical Note

4. TITLE AND SUBTITLE
Environmental Medicine Genome Bank (EMGB): Annual Report and Project Summary

5. FUNDING NUMBERS

6. AUTHOR(S)
N.J. Messinese, C.M. Lilly, L.A. Sonna

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Thermal & Mountain Medicine Division
U.S. Army Research Institute of Environmental Medicine
Kansas Street
Natick, MA 01760-5007

8. PERFORMING ORGANIZATION REPORT NUMBER
TN03-7

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Same as #7 above.

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Environmental Medicine Genome Bank (EMGB) project is an ongoing effort to identify and characterize genes relevant to environmental illnesses and to human physical performance. To accomplish this, the EMGB banks DNA samples from human volunteers who have participated in environmental and human performance studies or material obtained under approved Brigham and Women's Hospital protocols that would otherwise have been discarded. The EMGB maintains a registry of this phenotypic information. The EMGB can be used to identify polymorphisms in genes that are potentially of interest to environmental medicine and to obtain an estimate of the frequency of these polymorphisms in young, healthy U.S. adults because of the ethnically diverse and geographically dispersed backgrounds of the donors. Additionally, this resource also serves as a valuable source of control material for genetic studies of human diseases, such as asthma. The project is performed as part of a cooperative research and development agreement (CRDA) with the Division of Pulmonary and Critical Care Medicine at Brighsm and Women's Hopsital.

B-lympohcytes immortalized using the Epstein Barr Virus (EBV) are incorporated into the EMGB in attempts to maintain stocks of genetic material that are characterized by phenotype. These samples include cells and DNA from asthmatics, as well as from characterized non-asthmatics.

This report provides updated information about the samples currently stored in the EMGB and summarizes the uses of this resource to date. It is intended as a reference document for researchers who wish to make use of this resource. This report fulfills the annual reporting requirement of CRDA number DAMD 17-00-0017.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Genetics; Environmental Medicine; Asthma; Exercise; Physical Performance; Phenotype; Gene Bank	21
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

TABLE OF CONTENTS

SECTION	PAGE
EXECUTIVE SUMMARY	1
INTRODUCTION	2
MATERIALS AND METHODS	3
Volunteers	3
DNA isolation and storage	3
Aliquots	4
Lymphoyte Immortalization Protocol	4
Procedures for insuring confidentiality and anonymity of data	5
RESULTS	7
Contributing studies and sample use	7
Demographic information	13
DISCUSSION	15
REFERENCES	16
APPENDIX	18

EXECUTIVE SUMMARY

The Environmental Medicine Genome Bank (EMGB) project is an ongoing effort to identify and characterize genes relevant to environmental illnesses and to human physical performance. To accomplish this, the EMGB banks DNA samples from human volunteers who have participated in environmental and human performance studies or material obtained under approved Brigham and Women's Hospital protocols that would otherwise have been discarded. The EMGB maintains a registry of this phenotypic information. The EMGB can be used to identify polymorphisms in genes that are potentially of interest to environmental medicine and to obtain an estimate of the frequency of these polymorphisms in young, healthy U.S. adults because of the ethnically diverse and geographically dispersed backgrounds of the donors. Additionally, this resource also serves as a valuable source of control material for genetic studies of human diseases, such as asthma. The project is performed as part of a cooperative research and development agreement (CRDA) with the Division of Pulmonary and Critical Care Medicine at Brigham and Women's Hospital.

B-lymphocytes immortalized using the Epstein Barr Virus (EBV) are incorporated into the EMGB in attempts to maintain stocks of genetic material that are characterized by phenotype. These samples include cells and DNA from asthmatics, as well as from characterized non-asthmatics.

This report provides updated information about the samples currently stored in the EMGB and summarizes the uses of this resource to date. It is intended as a reference document for researchers who wish to make use of this resource. This report fulfills the annual reporting requirement of CRDA number DAMD 17-00-0017.

INTRODUCTION

Based on previous reports, it seems likely that there is a significant genetic contribution to some aspects of human physical performance (1-5;7;9;13) and to the susceptibility to environmental illness and injury. However, very few candidate genes have been identified, in part because few laboratories have access to large populations of well-characterized subjects drawn from a wide variety of genetic backgrounds. The U.S. Army Research Institute of Environmental Medicine (USARIEM) is uniquely qualified to undertake a search for these genes, by virtue of its access to Army personnel and its ability to define precisely those phenotypes relevant to environmental illnesses and human performance.

Large numbers of samples are typically needed to identify genes that contribute to complex traits (2). Accordingly, the Environmental Medicine Genome Bank (EMGB) banks DNA samples obtained from donor white blood cells and catalogues phenotypic information obtained over the course of multiple approved studies. By pooling samples and data from several studies, it becomes possible to undertake genetic analyses that would otherwise not be feasible.

The EMGB serves as an Institute resource, and anonymous aliquots from the bank are available to individual investigators upon request. This document summarizes the current contents of the bank.

A large fraction of the EMGB samples were obtained directly from terminally differentiated cell lines (donated peripheral blood mononuclear cells), and as such, represent non-renewable sources of genetic material. Immortalized lymphocytes have been added to the EMGB, making it possible to grow cells whenever necessary to extract more DNA. These cell lines represent a renewable source of genetic material. The samples of the EMGB that have not been immortalized will eventually be irreplaceably consumed.

MATERIALS AND METHODS

VOLUNTEERS

All subjects gave consent in accordance with 45 CFR 46. In study #1, subjects were recruited directly for the purpose of creating a core cohort of DNA samples from anonymous volunteers. Subjects from studies 2-6 were also participants in other USARIEM studies of environmental medicine and physical performance (Table 1). In study #7, the 56 immortalized samples were obtained through approved Brigham and Women's Hospital studies using material that otherwise would have been discarded. These samples are classified by asthmatic phenotypes such as FEV1, %FEV and skin testing for atopy. Subjects from studies # 8 and # 9 were recruited in order increase the number minorities and asthmatics, respectively, within the EMGB. History of asthma was documented and the FEV1 and FVC tests were performed the day of the blood draw. Study #10 includes subjects with chronic neuropathic pain after spinal cord injury. Phenotypic information obtained in study #10 does not include pulmonary function tests.

From each volunteer, 20 ml of blood were drawn into 10 ml, heparin-containing tubes. Samples drawn from locations remote from the analytical laboratory were shipped overnight (at room temperature) to the laboratory for processing.

DNA ISOLATION AND STORAGE

DNA is obtained from leukocyte nuclei after erythrocyte lysis, using the QIAamp Maxi Kit (Qiagen, Inc., Santa Clara, CA). The isolated DNA is stored in aqueous solution (in water), at a concentration of 35-150 ng / μl as determined by UV absorption at 260 nm.

The EMGB includes both samples obtained from white blood cells (which are not renewable) and immortalized B lymphocytes. The DNA from immortalized B lymphocytes were purified without nuclei isolation and therefore may include mitochondrial DNA. The DNA from these samples have been standardized to 50 ng / μl dilutions and tested using CLONTECH's β - Actin PCR primers (Cat# 5402), as outlined in the manufacture's instructions. This is performed to ensure constant quality of the template DNA during genetic studies. All future samples will meet this standard before aliquots are produced or dispensed.

ALIQUOTS

The EMGB was replenished by diluting the master DNA samples to 50 ng / uL, unless otherwise labeled, and dispensed in 25 uL aliquots. To minimize damage from repeated freeze-thaw cycles, each sample is divided into a master sample and several aliquots at the time of original isolation. At present, all samples are maintained at -80°C. Aliquots are used until exhausted. The master samples are thawed only when new aliquots are needed.

LYMPHOCYTE IMMORTALIZATION PROTOCOL

Immortalization Solution

The Epstein Barr Virus (EBV) Infection Solution is made from EBV-infected marmoset leukocyte cells. Marmoset cells produce viable virus, which is shed into the cell supernatant of the growth media (8). The supernatant is considered to be potent after the cells have been starved for approximately 1 week and when the phenol red pH indicator in the media has turned yellow. The cell suspension is collected in a 50 ml centrifuge tube and centrifuged at 400 x g for 10 min. The supernatant is filtered with a 0.2 μm filter to ensure the removal of marmoset cells. This virus containing filtrate solution is the infection solution that can be stored at 4°C for up to 3 months, or can be frozen at -80°C for long-term storage.

PBMC Purification

Subject peripheral blood monocytes (PBMC) are purified for immortalization. 20 ml of room temperature, heparin treated blood is added to 20 ml of Hank's balanced salt solution (HBSS) in a 50 ml conical centrifuge tube. Contents are gently mixed to make a homogenous suspension. Fifteen (15) mL of Histopaque – 1077 (Sigma cat# 10771) is added to two separate 50 mL Falcon centrifuge tubes. Histopaque is a density gradient solution that separates blood into two cell types: red blood cells (RBC) and PBMC. The blood-HBSS solution is slowly layered on top of each Histopaque solution. A sharp interface should be seen between the two solutions. The tubes are centrifuged at 400 x g for 30 min with the brake left off. Centrifugation results in four distinct density layers: plasma, PBMC, Histopaque 1077, and RBC. The second layer, the PBMC, is the buffy coat layer that contains the cells required to produce the immortalized cell lines. This layer is extracted and pipetted into a separate 50 mL centrifuge tube and washed with HBSS. Cells are then centrifuged at 400 x g for 10 min, the supernatant decanted, and the washing step repeated. Expected recovery is approximately $1x10^6$ cells per mL of blood drawn.

Immortilization of PBMCs

The PBMC are suspended in 2 ml of EBV Infection Solution and the vented tube is placed in the incubator at 37°C for 1-2 hours. 2 ml of media (RPMI w/ L-Glutamine, Penn-Strep, and 20% heat-inactivated FBS) is added to the cell suspension. 1 ml aliquots are added into four wells of a treated 24-well culture plate. Cell are placed at 37°C in 5% CO_2 for 4 days and transferred to T-25 or T-75 culture flasks at a concentration of $0.5x10^6$ cells per milliliter.

The immortalized cell lines are stored in liquid nitrogen in a step-wise fashion at concentrations of 5×10^6 cells per milliliter in DMSO Cell Freezing Media by Sigma (Catalog # C-6164). Cells are first stored at -20°C for 2 hours, then -80°C for 12 hours before finally transferring the cells to liquid nitrogen.

PROCEDURES FOR INSURING CONFIDENTIALITY AND ANONYMITY OF DATA

There are several layers of protection of confidentiality: First, the only temporary link between an individual and his EMGB number is retained in hard copy format in a locked cabinet. No electronic links between individual identifiers and EMGB numbers are retained. Second, the hard-copy data are ordered by study and EMGB number, not by last name Third, the EMGB database is on a free-standing (non-networked), password-protected computer. Fourth, many of the genotypes in the EMGB are themselves in shorthand notation, with the actual sequences they represent kept in laboratory notebooks or other files. Fifth, once no longer needed for purposes of data validation, the individual identifiers on the data sheets are defaced so as to permanently destroy the 'hard copy' link between individuals and their information.

The specific procedures for insuring confidentiality and anonymity are as follows:

a. Samples are collected at the study site and labeled according to the individual study guidelines. At the same time, the volunteer and PI fill out an EMGB data collection form (attached).
b. The samples and collection forms are shipped to the EMGB laboratory site. Upon arrival, each sample and its accompanying data form are assigned a unique EMGB number, which is written on the sample and on the form.
c. The forms are secured in a locked cabinet and constitute the only possible link between a subject and his or her genotypic information. The forms are retained to make sure original data are available to reconstruct or repair the electronic databases in the event of data corruption, loss, or mis-entry.
d. Data from the forms are stored electronically as follows:

i. All phenotypic data, except for names and social security numbers, are stored in an electronic database. These data are linkable to genotype data by means of the EMGB number.
ii. To prevent individuals from donating more than one sample to the EMGB, a separate electronic file containing nothing but names and social security numbers is maintained. This file is checked before any new samples are entered into the EMGB system. This file is NOT linked to EMGB numbers.

e. Genotype data is stored electronically and linked to phenotypic data through the EMGB numbers.
f. All electronic data is stored on a password-protected, non-networked computer. Backup copies of the data are stored in a locked cabinet.
g. Once no longer needed for data validation, the individual identifiers on the data sheets are defaced (blacked out with a magic marker) so as to destroy any potential 'hard copy' link between individuals and their EMGB information.

RESULTS

CONTRIBUTING STUDIES AND SAMPLE USE

Studies that have contributed samples to the EMGB and the current inventory of samples are listed in Table 1. This table also lists some of the phenotypic information available and summarizes some of the genotypic information that has been obtained on the samples to date. To date, 10 studies have contributed samples to the EMGB. DNA was obtained from most (but not all) donated samples, and some samples (especially those with low DNA yields) have been used in their entirety. At present, 431 different subjects have contributed to the EMGB.

Through the use of the EMGB during the past three years, the total DNA within the bank has diminished. Eighteen master aliquots from the original 283 terminal samples have been consumed. The changes in the total number of available samples during the past years are presented in Table 2. Sixteen samples have been consumed and are no longer available for study. Twenty samples have 5 or less aliquot available for study. Sixty samples were added to the EMGB in 2001 with immortalized B-lymphocytes. These represent a renewable supply of DNA that will not be completely consumed as long as the cell lines are maintained. Since the 2002 report, about 88 new samples have been added to the EMGB.

Table 1. Summary of the Contents of the EMGB.

Study #	Study Designation	PI, Division	Study Location	Study Dates	Samples Submitted	Samples Currently Banked	Phenotypic Information	Genotypes Studied
1	Normal Controls	Sonna, TMD	USARIEM	Mar - Apr 1998	62 Terminal	56	1,2	ACE, EOT,Gal-3, MCP-4, IL-18, CatS, NOS3
2	H98-07 Physical Fitness of Soldiers Entering and Completing Basic Combat Training and its Role in Injury Incidence	Sharp, MPD	Ft. Jackson, South Carolina	Jun - Jul 1998	152 Terminal	142	1,5,6,7,9,10	ACE, NOS1, EOT, Gal-3, MCP-4, IL-18, CatS, NOS3
3	H97-10 Warfighter Physiologic Status Monitoring: Body Core Temperature, Blood Oxygen Saturation and Environmental Symptoms during an Expedition to Mt. Logan, Canada	Muza, TMD	Mt. Logan, Canada	May - Jun 1999	13 Terminal	13	1,2,3,4,9	ACE, EOT, Gal-3, MCP-4, IL-18, CatS, NOS3
4	H98-09 Effect of Residence at Low and Moderate Altitudes on Arterial Oxygen Saturation at Moderate-to-High Altitudes	Muza, TMD	Pike's Peak, Colorado	Jun 1999	40 Terminal	40	1,2,4,8,9	EOT, Gal-3, MCP-4, IL-18, CatS, NOS3
5	H99-12/A-9212 Role of Exercise During Intermittent Exposures to Hypobaric Hypoxia on Acclimation to 4300 m	Beidleman, TMD	USARIEM	Oct 1999	7 Terminal	7*	1,2,4	EOT, Gal-3, MCP-4, IL-18, CatS, NOS3

6	H99-03 Role of Leukotrienes in High Altitude Illness	Muza, TMD	USARIEM	Jan - Feb 2000	9 Terminal 4 Immortal	13*	1,2,9	Gal-3, MCP-4, IL-18, CatS, NOS3
7	5-LO Immortalized Lymphocytes of Phenotyped Asthmatics and Non-Asthmatics.	Lilly, Pulmonary	BWH	April 2001	56 Immortal	56	1,3,4	C3a
8	Blood Collection from Normal Individuals	Lilly, Pulmonary	BWH	June 2001	26 Terminal	26	1-4,9	
9	Genetic Variants in 5-LO Product Production	Lilly, Pulmonary	BWH	Oct 2002 - Present	39 Terminal	39	1-4,9	
10	Chronic Neuropathic Pain after Spinal Chord Injury	Sang	MGH	Aug 2002- Present	23 Terminal	23	1,2,9	

*(One subject participated in both studies and is only counted once for purposes of the EMGB.)

Table 1 (Continued)

Key to Available Phenotypic Information

1. Age, race and gender
2. Smoking status
3. History of asthma or exercise-induced bronchospasm
4. Spirometry data
5. Spirometry before and after exercise
6. Army Physical Fitness Test scores
7. Army Physical Fitness Test scores before and after basic training
8. Oxygen saturation with increasing altitude
9. Height and weight

Key to Genotypes

ACE: Angiotensin Converting Enzyme Insertion/Deletion Polymorphism, Intron 16
NOS1: Neuronal Nitric Oxide Synthase CA repeat Polymorphism, Exon 29
EOT: Eotaxin 23ALA → 23 THR Polymorphism
MCP-4: Monocyte Chemo attractant Protein 4, Chromosome 17q11.2 promoter mutation
Gal-3: Galectin 3, Chromosome 14, Exons 3 and 6
CatS: Cathepsin S, Chromosome 1
NOS3: Nitric Oxide Synthase 3, Chromosome 7
IL-18: Interleukin 18, Promoter Region, Chromosome 2q12
C3a: Complement Cascade Protein

Table 2. EMGB Sample Availability

Year	2000	2001	2002	2003
Total Samples Submitted	283	343	343	431
Terminal Samples	283	283	283	371
Immortalized Samples	0	60	60	60
Samples Exhausted to Date	0	15	16	16
Totals Samples Available	**283**	**328**	**327**	**415**

Nine peer-reviewed manuscripts using samples from the EMGB have been accepted for publication over the past four years and a tenth is in preparation (Table 3). In the year 2000, the Angiotensin Converting Enzyme (ACE), nitric oxide synthase CA (NOS1), and the eotaxin (EOT) genotypes have been tested for population studies using the EMGB. The Angiotensin Converting Enzyme I/D polymorphism in intron 16 has been implicated by some as a marker of physical performance (5;7;9), though others have questioned this association (11;12). We concluded that the ACE genotype does not have a strong effect on aerobic power or muscular endurance in healthy, young American adults drawn from an ethnically and geographically diverse population (14). The neuronal NOS1 CA repeat polymorphism in exon 29 has been found to be associated with asthma, in a study to which the EMGB contributed samples (6). The EMGB was also used to determine the frequency of a novel mutation in the eotaxin gene that limits eotaxin secretion (10).

Two manuscripts that were supported by the EMGB were published in 2002 (Table III). The first analyzed sequence variants of the high-affinity IgE receptor, FcepsilonRI alpha gene that relate to the plasma level of total IgE. DNA from asthmatics and healthy subjects were screened for mutations by using single-stranded conformational polymorphism analysis. Three SNP's in the promoter region of the alpha chain gene were associated the level of IgE. The second manuscript reported variants in the epithelium-specific ETS-2 and ETS-3 genes that encode transcription factors, which drive genes known to be important in asthma pathogenesis. Three novel noncoding SNPs were found in the ETS-2 gene and four SNPs were found in the noncoding regions of the ETS-3 gene although none of the SNPs showed association to asthma susceptibility or asthma related phenotypes

Three manuscripts have been accepted for publication in 2003. Two studies investigated the association of MCP-4 and asthma. In the first study, plasma MCP-4 levels were associated with chronic-stable asthma and identified as a predictor of asthma diagnosis. In the second study, the role of MCP-4 gene sequence variants for determining plasma MCP-4 levels were investigated. Two SNPs were identified in the core promoter region preceding the transcription initiation site, one of which alters a consensus binding motif for the transcription factor YY-1. YY-1 had greater affinity for the more common (wild-type) sequence in an EMSA system and the wild type promoter had greater activity than the variant form in a reporter assay. Subjects bearing the less YY-1 activated variant form of MCP-4 had lower plasma MCP-4 levels than wild type bearing subjects. This study defines genetically determined decreased transcriptional activation as one mechanism that influences MCP-4 expression. The third study examined variants in the nitric oxide synthase genes and found that these correlated with nitric oxide levels in the exhaled air in asthmatic subjects. Strong associations were determined between a known functional NOS missense sequence variant in the endothelial nitric oxide gene and exhaled levels of nitric oxide.

The increasing number of peer-reviewed publications supported by the EMGB is correlated with the increase in samples available for study.

Table 3. Publications to date that have used the EMGB
1. Grasemann, H., C. Yandava, K. Storm van's Gravesande, A. Deykin, A. Pillari, J. Ma, L. Sonna, C. Lilly, M. Stampfer, E. Israel, E. Silverman, and J. Drazen. A neuronal NO synthase (NOS1) gene polymorphism is associated with asthma. Biochem.Biophys.Res.Commun. 272: 391-394, 2000.
2. Nakamura, H., A. Luster, T. Nakamura, K. Ho, L. Sonna, A. Deykin, E. Israel, J. Drazen, and C. Lilly. Variant eotaxin: Its effects on the asthma phenotype. J.Allergy Clin.Immunol. 108(6): 946-53, 2001
3. Sonna, L. A., M. A. Sharp, J. J. Knapik, M. Cullivan, K. C. Angel, J. F. Patton, and C. M. Lilly. Angiotensin-converting enzyme genotype and physical performance during US Army basic training. J.Appl.Physiol 91: 1355-1363, 2001.
4. Dikeidek, M., A. Bastille, L.A. Sonna, and C.M. Lilly. Environmental Medicine Genome Bank (EMGB): Hardy-Weinberg equilibrium at an eotaxin locus on chromosome 17. Technical Report T-02/14. USARIEM, Natick, MA, 2002.
5. Shikanai T, Silverman ES, Morse BW, Lilly CM, Inoue H, Drazen JM. Sequence variants in the Fc epsilon RI alpha chain gene. J Appl Physiol. 2002 Jul;93(1):37-41.
6. Baron RM, Palmer LJ, Tantisira K, Gabriel S, Sonna LA, Le L, Hallock A, Liberman TA, Drazen JM, Weiss ST, Silverman ES. DNA sequence variants in epithelium-specific ETS-2 and ETS-3 are not associated with asthma. Am J Respir Crit Care Med. 2002 Oct 1; 166(7):927-32.
7. Storm Van's Gravesande K, Wechsler ME, Grasemann H, Silverman ES, Le L, Palmer LJ, Drazen JM. The association of a missense mutation in the NOS3 gene with exhaled nitric oxide levels. Am J Respir Crit Care Med. 2003 Jul 15;168(2):228-31.
8. Kalayci O, Birben E, Wu L, Oguma T, Storm Van's Gravesande K, Subramaniam V, Sheldon HK, Silverman ES, Lilly CM. MCP-4 (CCL-13) core promoter genetic variants: Influence on YY-1 affinity and plasma levels. Am J Respir Cell Mol. Biol. 2003; in press (Epublication Jun 12).
9. Kalayci O, Sonna LA, Woodruff PG, Camargo CA, Luster AD, Lilly CM. Monocyte chemotactic protein-4 (MCP-4; CCL-13): A biomarker of asthma. J. Asthma 2003; in press.
10. Silverman E, Palmer LJ, Subramaniam A, Hallock A, Mathew S, Vallone J, Faffe DA, Shinanai T, Raby BB, Weiss ST, Stone SA. The transforming growth factor B1 promoter polymorphism C-508T is associated with asthma. For submission to : Am J Resp Crit Care Med.

DEMOGRAPHIC INFORMATION

A summary of the age, race and gender of the subjects for whom current samples exist in the EMGB is given in Table 4.

Table 4. Demographic Characteristics of the EMGB.
This table only includes sample numbers that still have DNA available for further experimentation. Full demographic data are not available for all samples.

Demographic	N =	%
Gender		
Male	218	57%
Female	167	43%
Ethnic Origin		
Asian	11	3%
African	75	21%
Caucasian	243	67%
Hispanic	27	8%
Native American	3	<1%
Other	1	<1%

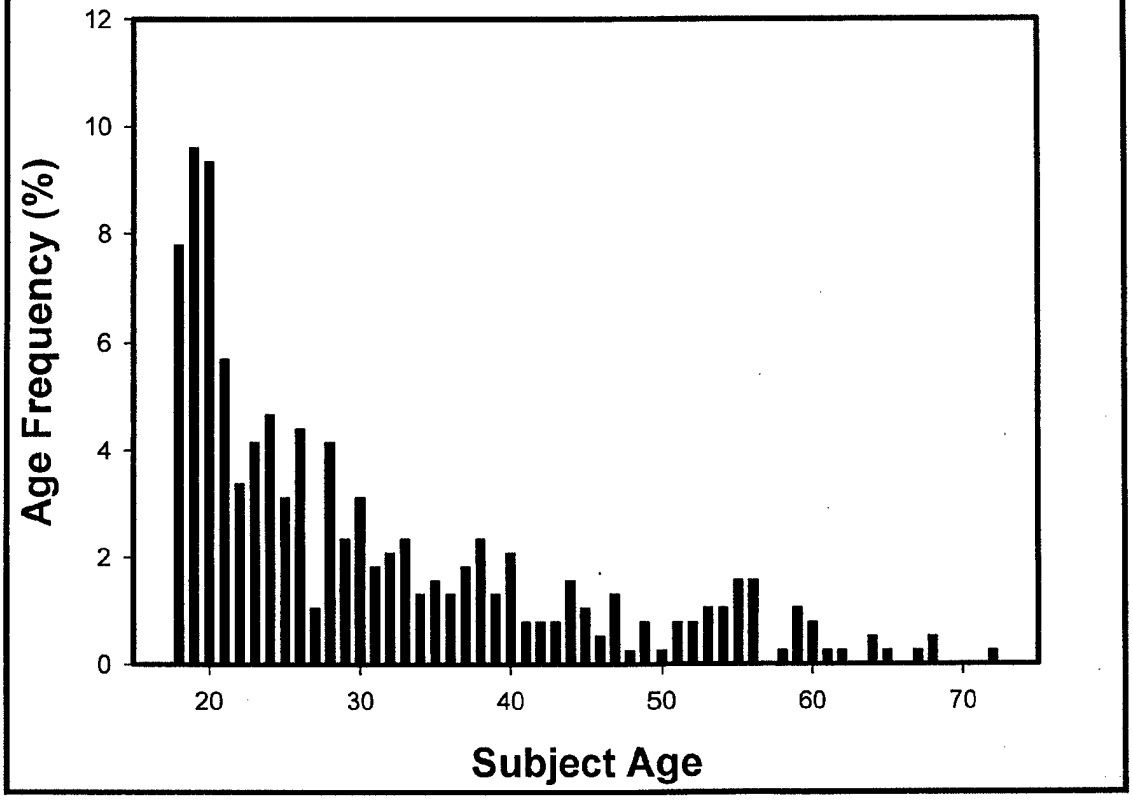
N = 416	Median	Interquartile Range
Age	26	20-37

The distribution of the ages of the donors at the time of sample collection is illustrated in Figure 1. The median age of the subjects who have donated to the EMGB is 26 (interquartile range, 20-38). Slightly more than half (57%) of the subjects are male. Subjects of ethnic backgrounds other than Caucasian donated about a third of the samples; subjects of African-American origin donated 21% of the samples. Site of origin are known for 65% of the subjects and include 44 different U.S. states, two U.S. territories, and three foreign nations.

Figure 1. EMGB Age Distribution

Age of subjects at time of donation to the EMGB

The histogram only includes subjects whose samples are currently available.



DISCUSSION

The EMGB consists of DNA samples obtained from an ethnically diverse and geographically dispersed population of subjects. This diversity makes the EMGB a valuable resource for several types of genetic studies. At present, we envision three principle uses for the bank. First, given a gene known or suspected to be of interest to environmental medicine, the EMGB can be used to identify new polymorphisms in this gene and to obtain an estimate of the frequency of these polymorphisms in young, healthy U.S. adults. Because the information collected in the EMGB includes both ethnic origin and gender, it is also possible to compare allele frequencies across important demographic subgroups. Second, the EMGB is a source of control material for genetic studies of human diseases, such as asthma. Third, some of the donor phenotypes in the EMGB (particularly those from study #2) are characterized well enough to allow genetic association studies.

Previously, one significant limitation of the EMGB was that samples were not easily renewed. Now that techniques have been adopted to immortalize lymphocytes, anonymous samples from individual donors can provide a renewable source of DNA.

The growth of he EMGB has been directly associated with an increase in the number of peer-reviewed publications that it has supported.

In summary, the current heterogeneity of the EMGB makes it a valuable resource of anonymous samples for genetic research. It has already proven to be of value in collaborative studies of human disease (6;10) and used to examine the genetic basis of physical performance (14). The addition of technology to allow source DNA to be renewed has greatly enhanced the longevity and potential value of this resource.

REFERENCES

1. Bouchard, C., F. Dionne, J. Simoneau, and M. Boulay. Genetics of aerobic performance. *Exerc.Sport Sci.Rev.* 20: 27-58, 1992.

2. Bouchard, C., R. Malina, and L. Perusse. Genetics of fitness and physical performance. Champaign, IL, Human Kinetics. 1997.

3. Bouchard, C., T. Rankinen, Y. Chagnon, T. Rice, L. Perusse, J. Gagnon, I. Borecki, P. An, A. Leon, J. Skinner, J. Wilmore, M. Province, and D. Rao. Genomic scan for maximal oxygen uptake and its response to training in the HERITAGE family study. *J.Appl.Physiol.* 88: 551-559, 2000.

4. Bray, M. Genomics, genes, and environmental interaction: The role of exercise. *J.Appl.Physiol.* 88: 788-792, 2000.

5. Gayagay, G., B. Yu, B. Hambly, T. Boston, A. Hahn, D. Celermajer, and R. Trent. Elite endurance athletes and the ACE I allele - the role of genes in athletic performance. *Hum.Genet.* 103: 48-50, 1998.

6. Grasemann, H., C. Yandava, K. Storm van's Gravesande, A. Deykin, A. Pillari, J. Ma, L. Sonna, C. Lilly, M. Stampfer, E. Israel, E. Silverman, and J. Drazen. A neuronal NO synthase (NOS1) gene polymorphism is associated with asthma. *Biochem.Biophys.Res.Commun.* 272: 391-394, 2000.

7. Hagberg, J., R. Ferrell, S. McCole, K. Wilund, and G. Moore. VO_{2max} is associated with ACE genotype in postmenopausal women. *J.Appl.Physiol.* 85: 1842-1846, 1998.

8. Miller, G. and M. Lipman. Release of infectious Epstein-Barr virus by transformed marmoset leukocytes. *Proc.Natl.Acad.Sci.U.S.A* 70: 190-194, 1973.

9. Montgomery, H., R. Marshall, H. Hemingway, S. Myerson, P. Clarkson, C. Dollery, M. Hayward, D. Holliman, M. Jubb, M. World, E. Thomas, A. Brynes, N. Saeed, M. Barnard, J. Bell, K. Prasad, M. Rayson, P. Talmud, and S. Humphries. Human gene for physical performance. *Nature* 393: 221-222, 1998.

10. Nakamura, H., A. D. Luster, T. Nakamura, K. H. In, L. A. Sonna, A. Deykin, E. Israel, J. M. Drazen, and C. M. Lilly. Variant eotaxin: its effects on the asthma phenotype. *J.Allergy Clin.Immunol.* 108: 946-953, 2001.

11. Rankinen, T., L. Perusse, J. Gagnon, Y. Chagnon, A. Leon, J. Skinner, J. Wilmore, D. Rao, and C. Bouchard. Angiotensin-converting enzyme ID polymorphism and fitness phenotype in the HERITAGE family study. *J.Appl.Physiol.* 88: 1029-1035, 2000.

12. Rankinen, T., B. Wolfarth, J. Simoneau, D. Maier-Lenz, R. Rauramaa, M. Rivera, M. Boulay, Y. Chagnon, L. Perusse, J. Keul, and C. Bouchard. No association between the angiotensin-converting enzyme ID polymorphism and elite endurance athlete status. *J.Appl.Physiol.* 88: 1571-1575, 2000.

13. Rankinen, T., L. Perusse, R. Rauramaa, M. A. Rivera, B. Wolfarth, and C. Bouchard. The human gene map for performance and health-related fitness phenotypes. *Med.Sci.Sports Exerc.* 33: 855-867, 2001.

14. Sonna, L. A., M. A. Sharp, J. J. Knapik, M. Cullivan, K. C. Angel, J. F. Patton, and C. M. Lilly. Angiotensin-converting enzyme genotype and physical performance during US Army basic training. *J.Appl.Physiol* 91: 1355-1363, 2001.

APPENDIX: EMGB Phenotypic Data Collection Form

This box is for Investigator use only	EMGB#:	Subject #

SUBJECT INFORMATION: Please answer all questions

Today's date:_____

Last Name: _____ Height:_____

First Name: _____ Weight: _____

Middle Initial: _____

SSN: _____ (Optional)

Gender (Check one): ____ MALE ____ FEMALE

Date of Birth: _____

Age: _____

Ethnic background (check one):	TOBACCO USE:
_____ Caucasian / White	_____ No previous tobacco habit
_____ African / Black	_____ Current smoker
_____ Asian / Pacific Islander	_____ Packs per day
_____ Hispanic / Latin American	_____ Number of years of smoking
_____ Native American	_____ Former smoker
_____ Other (please specify):	_____ Packs per day
	_____ Number of years of smoking
Place of birth (City/State)	_____ What year did you quit?
	_____ Cigars or Pipes (Years _____)
	_____ Chewing tobacco (Years _____)

Spirometry Date:

	#1	#2	#3	#4	#5	#6	#7	#8
FEV1:								
FVC:								